

Mid Term Part 3- Map Reduce
מספרי ת"ז של המגישים: 205614845,315313577

1) א.

Key:

$key = (i, k)$ כאשר i הוא האינדקס ההתחלתי לשורה i ו- k הוא האינדקס הסופי לשורה k .

כאשר הקלט שלנו הוא (a_i, \dots, a_k) עמודות i - k עוקבות במטריצה A .

values:

$value = (a_{ij} * b_j, \dots, a_{kj} * b_j)$ כאשר למעשה j הוא אינדקס העמודות כך ש $1 \leq j \leq n$ ו- b_j הוא האיבר בעמודה j – בוקטור b .

ב.

בשלב האיחוד, נאחד את כל הקואורדינטות של $A * b$ שקשורות ל i ול k :

עבור $key = (i, k)$ נרוץ על $1 \leq j \leq n$ כך שנקבל n ערכים לכל
 $tuple (a_{ij} * b_j, \dots, a_{kj} * b_j)$

כך שנקבל את הקבוצה הבאה:

$$(i, k; (a_{i1} * b_1, \dots, a_{k1} * b_1), \dots, (i, k; a_{in} * b_n, \dots, a_{kn} * b_n))$$

כך למעשה שלכל טאפל: $value$ מסעיף א' נוכל לקבל n קבוצות.

משמעות ה- key היא טאפל מהצורה (i, k) המסמנים כאמור את שורות i - k של מטריצה A .
כלומר זהו בלוק מהמטריצה בגודל $(k-i)$ שורות ו n עמודות.

משמעות ה $values$: $(a_{i1} * b_1, \dots, a_{k1} * b_1), \dots, (a_{in} * b_n, \dots, a_{kn} * b_n)$ (v_1, \dots, v_n)
כך ש- $v_j = a_{ij} * b_j, \dots, a_{kj} * b_j$ לכל $1 \leq j \leq n$. שזהו בדיוק הערך שקבענו במיפוי מסעיף א'.

נבצע reduce כך ש:

$$((i, k); v_1, \dots, v_n) = ((a_{i1} * b_1, \dots, a_{k1} * b_1), \dots, (a_{in} * b_n, \dots, a_{kn} * b_{1n})) \Rightarrow$$

$$((i, k); (\sum_{j=1}^n a_{ij} * b_j, \dots, \sum_{j=1}^n a_{kj} * b_j)) = (c_i, \dots, c_k)$$

כך שאלו ערכי התוצאה של הווקטור המתקבל מהמכפלה $A * b$ בשורות $i-k$.

כך שאם נפצל את המטריצה לקבוצות זרות לפי שורות רציפות מ $i-k$ ונאחד את כל התוצאות, נקבל לבסוף את התוצאה הנדרשת.

2. א.

Key:

$key = ((i, k), (j, m))$, הינו טאפל שמורכב משני טאפלים כאשר i הוא האינדקס ההתחלתי לשורה i , ו- k הוא האינדקס הסופי לשורה k .
כאשר הקלט שלנו הוא (a_i, \dots, a_k) עמודות $i-k$ עוקבות במטריצה הנתונה A .
כמו כן, j הוא האינדקס ההתחלתי לעמודה j ו- m הוא האינדקס הסופי לעמודה m כאשר (b_j, \dots, b_m) הן העמודות העוקבות מ- j במטריצה הנתונה B .

values:

$$: value = ((a_{iq} * b_{qj}, \dots, a_{iq} * b_{qm}), \dots, (a_{kq} * b_{qj}, \dots, a_{kq} * b_{qm}))$$

כאשר למעשה q הוא אינדקס השורות שרץ על שורות מטריצה B ועל עמודות מטריצה A .
למעשה זהו טאפל שמורכב מ- $i-k$ טאפלים כך שכל טאפל באורך $j-m$.

ב.

בשלב האיחוד, נאחד את כל הקואורדינטות של $A * B$ שקשורות ל $\text{key}: ((i,k),(j,m))$

עבור $\text{key} = \langle (i,k),(j,m) \rangle$ נרוץ על $1 \leq q \leq n$ כך שנקבל n ערכים לכל טאפל מהצורה:

$$\langle (i,k),(j,m) \rangle, ((a_{iq} * b_{qj}, \dots, a_{iq} * b_{qm}), \dots, (a_{kq} * b_{qj}, \dots, a_{kq} * b_{qm}))$$

כך למעשה נקבל:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} ((a_{i1} * b_{1j}, \dots, a_{i1} * b_{1m}), \dots, (a_{k1} * b_{1j}, \dots, a_{k1} * b_{1m})) \\ \vdots \\ ((a_{in} * b_{nj}, \dots, a_{in} * b_{nm}), \dots, (a_{kn} * b_{nj}, \dots, a_{kn} * b_{nm})) \end{pmatrix}$$

$$v_q = ((a_{iq} * b_{qj}, \dots, a_{iq} * b_{qm}), \dots, (a_{kq} * b_{qj}, \dots, a_{kq} * b_{qm})) \quad \text{כך ש:}$$

עבור מפתח $\text{key} = \langle (i,k),(j,m) \rangle$ לכל $1 \leq q \leq n$

כך שאנו מקבלים n ערכים לכל מפתח: $(\text{key}; v_1, \dots, v_n)$

משמעות ה- key היא טאפל מהצורה $\langle (i,k),(j,m) \rangle$ שמורכב משני טאפלים המסמנים בלוק של מטריצה A המורכב מהשורות i - k של המטריצה ובלוק של מטריצה B שמורכב מעמודות j - m של המטריצה. כך שכל מחשב מבצע חישוב של מכפלת המטריצה לפי שני הבלוקים הנ"ל.

משמעות ה- values :

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} ((a_{i1} * b_{1j}, \dots, a_{i1} * b_{1m}), \dots, (a_{k1} * b_{1j}, \dots, a_{k1} * b_{1m})) \\ \vdots \\ ((a_{in} * b_{nj}, \dots, a_{in} * b_{nm}), \dots, (a_{kn} * b_{nj}, \dots, a_{kn} * b_{nm})) \end{pmatrix}$$

$$v_q = ((a_{iq} * b_{qj}, \dots, a_{iq} * b_{qm}), \dots, (a_{kq} * b_{qj}, \dots, a_{kq} * b_{qm})) \quad \text{כך ש:}$$

עבור מפתח $\text{key} = \langle (i,k),(j,m) \rangle$ לכל $1 \leq q \leq n$ כך ש: $a_{pq} * b_{ql}$ זו מכפלה בין האיבר בשורה q במטריצה A (a_{pq}) לבין האיבר בשורה q ובעמוד l של מטריצה B (b_{ql})

נבצע reduce כך ש:

$$\left(key, \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right) = \left(((a_{i1} * b_{1j}, \dots, a_{i1} * b_{1m}), \dots, (a_{k1} * b_{1j}, \dots, a_{k1} * b_{1m}), \dots, ((a_{in} * b_{nj}, \dots, a_{in} * b_{nm}), \dots, (a_{kn} * b_{nj}, \dots, a_{kn} * b_{nm})) \right) \Rightarrow$$

$$\left(key = \langle (i, k), (j, m) \rangle; \begin{pmatrix} \sum_{q=1}^n a_{iq} * b_{qj} & \dots & \sum_{q=1}^n a_{iq} * b_{qm} \\ \vdots & \ddots & \vdots \\ \sum_{q=1}^n a_{kq} * b_{qj} & \dots & \sum_{q=1}^n a_{kq} * b_{qm} \end{pmatrix} \right) = \begin{pmatrix} c_{ij} & \dots & c_{im} \\ \vdots & \ddots & \vdots \\ c_{kj} & \dots & c_{km} \end{pmatrix}$$

כלומר פעולת ה reduce נותנת לנו לבסוף בלוק של מטריצת התוצאה C כאשר $A*B=C$ כך שמתקבלת תת מטריצה C עם $(k-i+1)$ שורות ו $(m-j+1)$ עמודות.

כך שאם נפצל את המטריצות לקבוצות זרות $\{(i, k), (j, m)\}$ שאיחודן ייתן לנו את כל הצירופים של העמודות והשורות של מטריצה C ונאחד את כל התוצאות, נקבל לבסוף את התוצאה הנדרשת C שהיא תוצאת המכפלה $A*B=C$.